Reachability Oracles for Directed Transmission Graphs*

Haim Kaplan[†] *Wolfgang Mulzer*[‡] *Liam Roditty*[§] *Paul Seiferth*[‡]

Abstract

Let $P \subset \mathbb{R}^d$ be a set of n points in the d dimensions such that each point $p \in P$ has an *associated radius* $r_p > 0$. The *transmission graph* G for P is the directed graph with vertex set P such that there is an edge from p to q if and only if $d(p,q) \leq r_p$, for any $p, q \in P$.

A *reachability oracle* is a data structure that decides for any two vertices $p, q \in G$ whether G has a path from p to q. The quality of the oracle is measured by the space requirement $S(n)$, the query time $Q(n)$, and the preprocessing time. For transmission graphs of one-dimensional point sets, we can construct in $O(n \log n)$ time an oracle with $Q(n) = O(1)$ and $S(n) = O(n)$. For planar point sets, the ratio Ψ between the largest and the smallest associated radius turns out to be an important parameter. We present three data structures whose quality depends on Ψ: the first works only for $\Psi < \sqrt{3}$ and achieves $Q(n) = O(1)$ with $S(n) = O(n)$ and preprocessing time $O(n \log n)$; the second data structure gives $Q(n) = O(\Psi^3\sqrt{n})$ and $S(n) = O(\Psi^5 n^{3/2})$; the third data structure is randomized with $Q(n) = O(n^{2/3} \log^{1/3} \Psi \log^{2/3} n)$ and $S(n) = O(n^{5/3} \log^{1/3} \Psi \log^{2/3} n)$ and answers queries correctly with high probability.

1 Introduction

Representing the connectivity of a graph in a space efficient, succinct manner, while supporting fast queries, is one of the most fundamental data structuring questions on graphs. For an undirected graph, it suffices to compute the connected components and to store with each vertex a label for the respective component. This leads to a linear-space data structure that can decide in constant time if any two given vertices are connected. For undirected graphs, however, connectivity is not a symmetric relation any more, and the problem turns out to be much more challenging. Thus, if G is a directed graph, we say that a vertex s can *reach* a vertex t if there is a directed path in G from s to t. Our goal is to construct a *reachability oracle*, a space efficient data structure that answers *reachability queries*, i.e., that determines for any pair of query vertices s and t whether s

*This work is supported in part by GIF project 1161 & DFG project MU/3501/1. A preliminary version appeared as Haim Kaplan, Wolfgang Mulzer, Liam Roditty, and Paul Seiferth. *Spanners and Reachability Oracles for Directed Transmission Graphs.* Proc. 31st SoCG, pp. 156–170.

[†]School of Computer Science, Tel Aviv University, Israel, haimk@post.tau.ac.il

[‡]Institut für Informatik, Freie Universität Berlin, Germany {mulzer,pseiferth}@inf.fu-berlin.de

[§]Department of Computer Science, Bar Ilan University, Israel liamr@macs.biu.ac.il

can reach t. The quality of a reachability oracle for a graph with n vertices is measured by three parameters: the *space* $S(n)$, the *query time* $Q(n)$ and the *preprocessing time*. The simplest solution stores for each pair of vertices whether they can reach each other, leading to a reachability oracle with $\Theta(n^2)$ space and constant query time. For sparse graphs with $O(n)$ edges, storing just the graph and performing a breadth first search for a query yields an $O(n)$ space oracle with $O(n)$ query time. Interestingly, other than that, we are not aware of any better solutions for general directed graphs, even sparse ones. Thus, any result that simultaneously achieves subquadratic space and sublinear query time would be of great interest. A lower bound by Pătraşcu [11] shows that we cannot hope for $o(\log n)$ query time with linear space, but it does not rule out constant time queries with slightly superlinear space. In the absence of progress towards non-trivial reachability oracles or better lower bounds, solutions for special cases become important. For directed planar graphs, after a long line of research [2, 3, 6, 7, 12], Holm, Rotenberg and Thorup presented a reachability with optimal parameters [8]. This result, as well as most other previous results, is actually not only a reachability oracle but it can also return the approximate shortest path distance between the query vertices.

Transmission graphs constitute a graph class that shares many similarities with planar graphs: let $P \subset \mathbb{R}^2$ be a set of points where each point $p \in P$ has a (transmission) radius r_p associated with it. The transmission graph has vertex set P and a *directed* edge between two distinct points $p, q \in P$ if and only if $|pq| \leq r_p$, where $|pq|$ denotes the Euclidean distance between p and q. Transmission graphs are a common model for directed sensor networks [9, 10, 13]. In this geometric context, it is natural to consider a more general type of query where the target point is an arbitrary point in the plane rather a vertex of the graph. In this case, a vertex $s \in P$ can reach a *point* $q \in \mathbb{R}^2$ if there is a *vertex* $t \in P$ such that s reaches t and such that $|tq| \leq r_t$. We call such queries *geometric reachability queries* and oracles for them *geometric reachability oracles*. To avoid ambiguities, we sometimes use the term *standard* reachability query/oracle when referring to the case where the query consists of two vertices.

Our Results. In Section 3 we will see that one-dimensional transmission graphs admit a rich structure that can be exploited to construct a simple linear space geometric reachability oracle with constant query time and $O(n \log n)$ preprocessing time.

In two dimensions, the situation is much more involved. Here, it turns out that the *radius ratio* Ψ, the ratio of the largest and the smallest transmission radius in P, is an important parameter. If Ψ is less than $\sqrt{3}$, we can turn the transmission graph into a planar graph in $O(n \log n)$ time, while preserving the reachability structure and keeping the number of vertices linear in n. As mentioned above, for planar graphs there is a linear time construction of a reachability oracle with linear space and constant query time [8]. This construction yields a standard reachability oracle. However, in a companion paper we show that any standard reachability oracle can be transformed into a geometric one by paying an *additive* overhead of $O(\log n \log \Psi)$ in the query time and of $O(n \log \Psi)$ in the space [9]. Our final construction needs $O(n)$ space and has query time $O(\log n)$ for geometric queries and $O(1)$ for standard queries. It can be found in Section 4.1.

When $\Psi \geq \sqrt{3}$, we do not know how to find a planar graph representing the reachability of G. Fortunately, we can use a theorem by Alber and Fiala that allows us to find a small and balanced separator with respect to the area of the union of the disks [1]. This leads to a standard reachability oracle with query time $O(\Psi^3\sqrt{n})$ and space and preprocessing time $O(\Psi^5 n^{3/2})$, see Section 4.2. When Ψ is even larger, we can use random sampling combined with a quadtree of logarithmic depth to obtain a standard reachability oracle with query time $O(n^{2/3}\log^{1/3}\Psi\log^{2/3}n)$, space $O(n^{5/3}\log^{1/3}\Psi\log^{2/3}n)$, and preprocessing time $O(n^{5/3}(\log\Psi+\log n)\log^{1/3}\Psi\log^{2/3}n)$. Refer to Section 4.3. Again, we can transform both oracles into geometric reachability oracles using the result from the companion paper [9]. Since the overhead is additive, the transformation does not affect the performance bounds.

2 Preliminaries and Notation

Unless stated otherwise, we let $P \subset \mathbb{R}^2$ denote a set of n points in the plane, and we assume that for each point p, we have an *associated radius* $r_p > 0$. Furthermore, we assume that the input is scaled so that the smallest associated radius is 1. The elements in P are called *vertices*. The *radius ratio* Ψ of P is defined as $\Psi = \max_{p,q\in P} r_p/r_q$. Given a point $p \in \mathbb{R}^2$ and a radius r, we denote by $D(p,r)$ the closed disk with center p and radius r. If $p \in P$, we use $D(p)$ as a shorthand for $D(p,r_p)$. We write $C(p,r)$ for the boundary circle of $D(p,r)$.

Our constructions for the two dimensional reachability oracles make extensive use of planar grids. For $i \in \{0,1,\dots\}$, we denote by $\mathcal{Q}_i$ the *grid at level* i. It consists of axis-parallel squares with diameter 2^i that partition the plane in grid-like fashion (the *cells*). Each grid $\mathcal{Q}_i$ is aligned so that the origin lies at the corner of a cell. We assume that our model of computation can find in constant time for any given point the grid cell that contains it.

In the one dimensional case, our construction immediately yields a geometric reachability oracle. In the two dimensional case, we are only able to construct standard reachability oracles directly. However, we can use the following result from the companion paper to transform these oracles into geometric reachability oracles in a black-box fashion [9].

Theorem 2.1 (Theorem 4.3 in [9])**.** *Let G be the transmission graph for set P of n points in the plane with radius ratio Ψ. Given a reachability oracle for G that uses $S(n)$ space and has query time $Q(n)$, we can compute in time $O(n\log n\log\Psi)$ a geometric reachability oracle with space $S(n)+O(n\log\Psi)$ and query time $Q(n)+O(\log n\log\Psi)$.*

To achieve a fast preproccesing time, we need a sparse approximation of the transmission graph G. Let $\varepsilon > 0$ be constant. A $(1+\varepsilon)$-*spanner* for G is a sparse subgraph $H \subseteq G$ such that for any pair of vertices p and q in G we have $d_H(p,q) \leq (1+\varepsilon)d_G(p,q)$ where d_H and d_G denote the shortest path distance in H and in G. The companion paper shows that $(1+\varepsilon)$-spanners for transmission graphs can be constructed quickly [9].

Theorem 2.2 (Theorem 3.12 in [9])**.** *Let G be the transmission graph for a set P of n points in the plane with radius ratio Ψ. For any fixed $\varepsilon > 0$, we can compute a $(1+\varepsilon)$-spanner for G with $O(n)$ edges in time $O(n(\log n+\log\Psi))$ using space $O(n\log\Psi)$.*

3 Reachability Oracles for 1-dimensional Transmission Graphs

In this section, we prove the existence of efficient reachability oracles for one-dimensional transmission graphs and show that they can be computed quickly.

Theorem 3.1. *Let G be the transmission graph of an n-point set $P \subset \mathbb{R}$. We can construct in $O(n \log n)$ time a geometric reachability oracle for G with space $S(n) = O(n)$ and query time $Q(n) = O(1)$.*

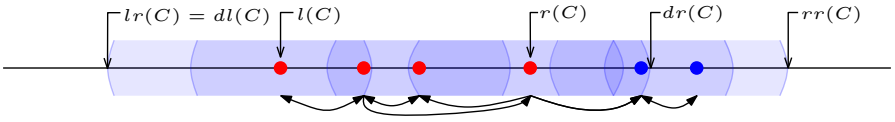

Fig. 1: A strongly connected component C (red) and its landmarks. The right reachpoint is defined by a vertex of another component (blue).

It suffices to consider the reachability for each *strongly connected components* (SCCs) of G. Let $\mathcal{C}$ be the set of the SCCs of G, and let $C \in \mathcal{C}$ be one of them. We say that the SCC C can *reach* a point $q \in \mathbb{R}^2$, if there is a vertex p in G with $q \in D(p)$ and directed path in G from a vertex in C to p. We say that C can reach an SCC $D \in \mathcal{C}$ if C can reach a vertex in D. By strong connectivity, this means that all vertices in C can reach all verticies in D. Next, we define several landmarks related to the SCC C, see Figure 1: the *leftmost point* of C, $\mathrm{l}(C)$, is the vertex in C with the smallest x-coordinate; the *left reachpoint* of C, $\mathrm{lr}(C)$, is the leftmost point in $\mathbb{R}$ that lies in a ball around a vertex in P reachable from C; and the *direct left reachpoint* of C, $\mathrm{dl}(C)$, is the leftmost point in $\mathbb{R}$ that lies in a ball around a vertex in C, i.e., $\mathrm{dl}(C) = \min_{p \in C}(p - r_p)$. The rightmost point $\mathrm{r}(C)$, right reachpoint $\mathrm{rr}(C)$, and the direct right reachpoint $\mathrm{dr}(C)$ are defined analogously. The *interval* of C, I_C, is defined as $I_C = [\mathrm{l}(C), \mathrm{r}(C)]$.

Lemma 3.2. *Let $C \in \mathcal{C}$ be a strongly connected component, and let $s \in C$ be a vertex in C. For any point $q \in \mathbb{R}^2$, the vertex s can reach q if and only if $q \in [\mathrm{lr}(C), \mathrm{rr}(C)]$.*

Proof. If s can reach q, then $q \in [\mathrm{lr}(C), \mathrm{rr}(C)]$, by the definition of $\mathrm{lr}(C)$ and $\mathrm{rr}(C)$. Conversely, let $q \in [\mathrm{lr}(C), \mathrm{rr}(C)]$, and assume w.l.o.g that q lies to the left of s. Let $p \in P$ be the vertex that defines the left reachpoint, i.e., $\mathrm{lr}(C) = p - r_p$. By definition of $\mathrm{lr}(C)$, there is a path $s = p_1 p_2 \dots p_k = p$ from s to p in G. Since G is a transmission graph, we have $|p_i - p_{i+1}| \leq r_{p_i}$, for $i = 1, \dots, k-1$, so the disks $D(p_i)$ cover the entire interval $[\mathrm{lr}(C), p]$. Thus, there is a p_i with q in $q \in D(p_i)$. By definition, G contains a path from s to p_i and hence s can reach q. □

Lemma 3.2 suggests the following reachability oracle with $O(n)$ space and $O(1)$ query time: for each SCC $C \in \mathcal{C}$, store the reachpoints $\mathrm{lr}(C)$ and $\mathrm{rr}(C)$; and for each vertex $p \in P$, store the SCC of G that contains it. Given a query p, q, where p is a vertex and q a point in $\mathbb{R}^2$, we look up the SCC C for p, and we return YES if and only if $q \in [\mathrm{lr}(C), \mathrm{rr}(C)]$.

The Structure of the Components. To compute the reachpoints efficiently, we must investigate the structure of the SCCs in one-dimensional transmission graphs more deeply.

Lemma 3.3. *The intervals* $\{I_C \mid C \in \mathcal{C}\}$ *for the SCCs form a* laminar set family, *i.e., for any two SCCs* $C, D \in \mathcal{C}$, *we have either* $I_C \cap I_D = \emptyset$, $I_C \subseteq I_D$, *or* $I_D \subseteq I_C$.

Proof. Fix two distinct SCCs $C, D \in \mathcal{C}$, and suppose that $I_C \cap I_D \neq \emptyset$, but neither $I_C \subseteq I_D$ nor $I_D \subseteq I_C$ (see Figure 2). Then one endpoint of I_C lies in I_D, and vice versa. Since C is strongly connected, for every $x \in I_C$, there is a vertex $p \in C$ with $|p - x| \leq r_p$. The same holds for D. Thus, since the endpoints of I_C and I_D lie in P, strong connectivity implies that C can reach D and that D can reach C. However, this means that $C = D$, although we assumed them to be distinct. □

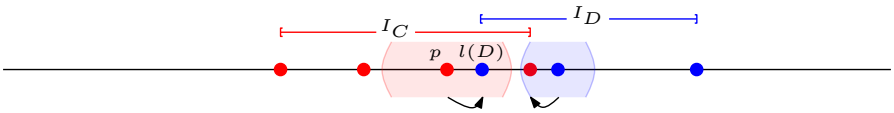

Fig. 2: Two SCCs C and D whose intervals overlap. There must be a vertex $p \in C$ whose disk contains $l(D)$ and vice versa, so C and D cannot be distinct.

By Lemma 3.3, we can obtain a rooted forest with vertex set $\mathcal{C}$ by making $C \in \mathcal{C}$ a child of $D \in \mathcal{C}$, if and only if $I_C \subset I_D$. If necessary, we add a common root node to get a rooted tree T. The next lemma characterizes the left and the right reachpoints.

Lemma 3.4. *Let* $C \in \mathcal{C}$ *be a SCC. The left reachpoint* $\mathrm{lr}(C)$ *of* C *is either* $\mathrm{dl}(C)$ *or* $\mathrm{dl}(D)$, *where* D *is a sibling of* C *in* T. *The situation for the right reachpoints is analogous.*

Proof. Let $\overline{C}$ be the parent of C in T. Since $I_C \subseteq I_{\overline{C}}$, and since the SCCs C and $\overline{C}$ are distinct, $\overline{C}$ can reach C, but C cannot reach $\overline{C}$. Furthermore, since the endpoints of $I_{\overline{C}}$ are vertices of $\overline{C}$, we cannot reach any SCC outside $I_{\overline{C}}$ from C, otherwise Lemma 3.2 would imply that C can reach also $\overline{C}$.

By the definition of (direct) left reachpoint, there is an SCC $D \in \mathcal{C}$ with $\mathrm{lr}(C) = \mathrm{dl}(D)$, possibly $D = C$. Since C cannot reach outside $I_{\overline{C}}$, it follows that $I_D \subseteq I_{\overline{C}}$ and that D is a descendant of $\overline{C}$. Assume that $D \neq C$ and that D is not a sibling of C. Then, by Lemma 3.3, there is a sibling D' of C with $I_D \subset I_{D'}$. Since $\mathrm{lr}(C) = \mathrm{dl}(D)$, we can go from C to D and, by Lemma 3.2, also to D'. Hence, $\mathrm{lr}(D') = \mathrm{lr}(C)$. Since $I_D \subset I_{D'}$ and $D \neq D'$, it follows that $\mathrm{l}(D') < \mathrm{dl}(D)$. Hence, we get $\mathrm{lr}(C) = \mathrm{lr}(D') \leq \mathrm{l}(D') < \mathrm{lr}(C)$, a contradiction. □

Reachability Between Siblings. By Lemma 3.4, for an SCC $C \in \mathcal{C}$, it suffices to search for $\mathrm{lr}(C)$ and $\mathrm{rr}(C)$ among the siblings of C in T. Let $C_1, \ldots, C_k$ be the children of a node in T, sorted from left to right according to their intervals. To compute the left reachpoints of $C_1, \ldots, C_k$, we proceed as follows: we set $\mathrm{lr}(C_1) = \mathrm{dl}(C_1)$, and we push C_1 onto an empty stack S. Then we go through $C_2, \ldots, C_k$, from left to right. For the current child C_i, we initalize the tentative left reachpoint $\mathrm{lr}(C_i) = \mathrm{dl}(C_i)$. While the

current tentative reachpoint lies to the left of the right interval endpoint for the top of the stack, we pop the stack, and we update the tentative reachpoint of C_i to the left reachpoint of the popped component, if that reachpoint lies further to the left. Then, we push C_i onto the stack and proceed to the next child; see Algorithm 1.

```
1 lr(C_1) ← dr(C_1)
2 push C_1 onto an empty stack S
3 for i ← 2 to k do
4 |   lr(C_i) ← dr(C_i)
5 |   while S ≠ ∅ and lr(C_i) ≤ r(top(S)) do
6 |   |   D ← pop(S)
7 |   |   lr(C_i) ← min{lr(C_i), lr(D)}
8 |   push C_i onto S
```

Algorithm 1: Computing left reachpoints.

The right reachpoints are computed analogously. In the next lemma, we prove that this procedure correctly computes the reachability among nodes of T in time $O(n \log n)$.

Lemma 3.5. *We can compute the reachability for all nodes in T in $O(n \log n)$ time.*

Proof. We apply Algorithm 1 for each set of siblings. The total time for sorting the intervals is $O(n \log n)$. The total time for computing the direct left reachpoints is linear, and each SCC is pushed and popped at most once, again taking linear time.

It remains to prove correctness. For this, consider the sorted siblings $C_1, \ldots, C_k$. During Algorithm 1, we maintain the following invariant: let C_i be the current component. Then, all components C_j, $j = 1, \ldots, i-1$, have the correct left reachpoint, and S contains precisely those components C_j, $j = 1, \ldots, i-1$, that cannot be reached by any component C_l with $j < l < i$. The invariant holds for C_1: if $\mathrm{dl}(C_1)$ were different from $\mathrm{lr}(C_1)$, there would be another component D with $\mathrm{dl}(D) = \mathrm{lr}(C_1)$. The component D cannot be to the left of C_1, as C_1 is the leftmost sibling, and it cannot be to the right of C_1, since then $I_{C_1} \subseteq [\mathrm{lr}(D), \mathrm{rr}(D)]$, and C and D would collapse into one SCC, by Lemma 3.2. Thus $\mathrm{dl}(C_1) = \mathrm{lr}(C_1)$.

For $i > 1$, let $p \in P$ be the vertex with $\mathrm{lr}(C_i) = p - r_p$. If $p \in C_i$, then $\mathrm{lr}(C_i) = \mathrm{dl}(C_i)$, and we are done. Otherwise, let π be a path from C_i to p, and let D be the first component that π visits after C_i. Then, D must be to the left of C_i (otherwise $C_i = D$) and $\mathrm{lr}(C_i) = \mathrm{lr}(D)$. First suppose that D is in S. Since $\mathrm{dl}(C_i)$ lies to the left of $r(C)$, the algorithm pops the stack until it reaches D. After that, it sets $\mathrm{lr}(C_i)$ to $\mathrm{lr}(D)$ and stops, by the invariant. Thus, $\mathrm{lr}(C_i)$ is set correctly and the invariant is maintained. If D is not in S, the invariant ensures that there is a component C_l on S that can reach D and that lies between C_i and D, i.e., $\mathrm{lr}(C_l) = \mathrm{lr}(D)$. As before, we see that the algorithm pops the stack until it discovers C_l and then correctly sets the left reachpoint of C_i while maintaining the invariant. □

To establish Theorem 3.1, it remains to describe how to find the SCCs efficiently without constructing the transmission graph explicitly.

Lemma 3.6. *The SCCs can be computed in $O(n \log n)$ time.*

Proof. Recall the Kosaraju-Sharir algorithm [4]: first, it performs a DFS of G and records the order in which the vertices are visited for the last time. Then it performs a second DFS in the *transpose graph* G' in which the directions of all edges have been reversed.

To implement this algorithm, we need two operations: given a vertex $p \in P$, find an unvisited vertex q such that pq is an edge of G or an edge of G'. For G, this can easily be done in $O(\log n)$ time: store the points of P in a balanced search tree. When a point p is visited for the first time, remove it from the tree. When looking for an outgoing edge from a vertex p, determine the predecessor and the successor of p in the current set, and check the distances. For G', we proceed similarly, but we use an *interval tree* to store the r_p-balls around the vertices in P [5]. When a vertex p is visited for the first time, we delete the corresponding r_p-ball from the interval tree. When we need to find an outgoing edge from a vertex p, we use the interval tree to find one ball that contains p. Again, this can be done in $O(\log n)$ time. □

4 Reachability Oracles for 2-dimensional Transmission Graphs

In the following sections we present three different geometric reachability oracles for transmission graphs in $\mathbb{R}^2$. By Theorem 2.1, we can focus on the construction of standard reachability oracles since they can be extended easily to geometric ones. This has no effect on the space and query time bounds, expect for the oracle given in Section 4.1. This oracle applies for $\Psi < \sqrt{3}$, it needs space $O(n \log n)$ and has query time $O(1)$. Thus, the transformation from standard reachability to geometric reachability increases the query time to $O(\log n)$.

4.1 Ψ is less than $\sqrt{3}$

Suppose that $\Psi \in [1, \sqrt{3})$. In this case, we show that we can make G planar by first removing unnecessary edges and then resolving edge crossings by adding $O(n)$ additional vertices. This will not change the reachability between the original vertices. The existence of efficient reachability oracles then follows from known results for directed planar graphs. The main goal is to prove the following lemma.

Lemma 4.1. *Let G be the transmission graph for a planar n-point set P with $\Psi < \sqrt{3}$. In $O(n \log n)$ time, we can find a plane graph $H = (V, E)$ such that*

(i) $|V| = O(n)$ and $|E| = O(n)$;

(ii) $P \subseteq V$; and

(iii) for any $p, q \in P$, p can reach q in G if and only if p can reach q in H.

Given Lemma 4.1, we can obtain our reachability oracle from known results.

Theorem 4.2. *Let G be the transmission graph for a two-dimensional set P of n points, and suppose the radius ratio Ψ is less than $\sqrt{3}$. Then, we can construct in $O(n \log n)$ time a standard reachability oracle for G with $S(n) = O(n)$ and $Q(n) = O(1)$ or a geometric reachability oracle for G with $S(n) = O(n)$ and $Q(n) = O(\log n)$.*

Proof. We apply Lemma 4.1 and construct the distance oracle of Holm, Rotenberg, and Thorup for the resulting graph [8]. This distance oracle can be constructed in linear time, it needs linear space, and it has constant query time. The result for the geometric reachability oracle follows from Theorem 2.1. □

We prove Lemma 4.1 in three steps. First, we show how to make G sparse without changing the reachability. Then, we show how to turn G into a planar graph. Finally, we argue that we can combine these two operations to get the desired result.

Obtaining a Sparse Graph. We construct a subgraph $H \subseteq G$ with the same reachability as G but with $O(n)$ edges and $O(n)$ edge crossings. The bounded number of crossings will allow us to obtain a planar graph later on. Consider the grid $\mathcal{Q}_0$, and let $\sigma \in \mathcal{Q}_0$ be a grid cell. We say that an edge of G *lies in* σ if both endpoints are contained in σ. The *neighborhood* $N(\sigma)$ of σ consists of the 7×7 block of cells in $\mathcal{Q}_0$ with σ at the center. Two grid cells are *neighboring* if they lie in each other's neighborhood. Since a cell in $\mathcal{Q}_0$ has side length $\sqrt{2}/2$, the two endpoints of every edge in G must lie in neighboring grid cells.

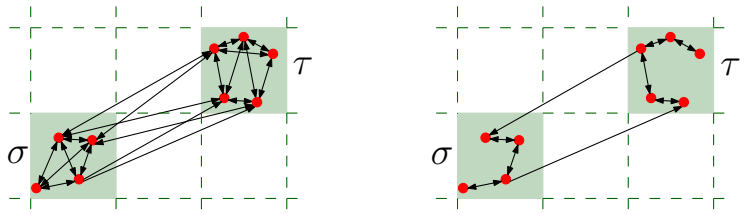

Fig. 3: The vertices and edges of two neighboring cells of G (left) and of H (right)

The subgraph H has vertex set P, and we pick the edges as follows (see also Figure 3): for each non-empty cell $\sigma \in \mathcal{Q}_0$, we set $P_\sigma = P \cap \sigma$, and we compute the Euclidean minimum spanning tree (EMST) T_σ of P_σ. For each edge pq of T_σ, we add the directed edges pq and qp to H. Then, for every cell $\tau \in N(\sigma)$, we check if there are any edges from σ to τ in G. If so, we add an arbitrary such edge to H. The following lemma states properties of H.

Lemma 4.3. *The graph H*

(i) has the same reachability as G;

(ii) has $O(n)$ edges;

(iii) can be constructed in $O(n \log n)$ time; and

(iv) the straight line embedding of H in the plane with vertex set P has $O(n)$ edge crossings.

Proof. (i): All edges of H are also edges of G: inside a non-empty cell σ, P_σ induces a clique in G, and the edges of H between cells lie in G by construction. It follows that H does not increase the reachability. Now let pq be an edge in G. We show that there is a path from p to q in H: if pq lies in a cell σ of $\mathcal{Q}_0$, we take the path along the EMST T_σ. If pq goes from a cell σ to another cell τ, then there is an edge uv from σ to τ in H, and we take the path in T_σ from p to u, then the edge uv, and finally the path in T_τ from v to q.

(ii): For a nonempty cell σ, we create $|P_\sigma| - 1$ edges inside σ. Furthermore, since $|N(\sigma)|$ is constant, there are at most $O(1)$ edges between σ and other cells. Thus, H has $O(n)$ edges.

(iii): Since we assumed that we can find the cell for a vertex $p \in P$ in constant time, we can easily compute the sets P_σ, $\sigma \in \mathcal{Q}_0$ nonempty, in time $O(n \log n)$ Computing the EMST T_σ for a cell σ needs time $O(|P_\sigma| \log |P_\sigma|)$, for a total of $O(n \log n)$. To find the edges between neighboring cells, we build a Voronoi diagram together with a point location structure for each set P_σ. Again, this takes $O(n \log n)$ total time. Let σ and τ be two neighboring cells. For each point in P_σ, we locate the nearest neighbor in P_τ using the Voronoi diagram. If there is a point $p \in P_\sigma$ whose nearest neighbor $q \in P_\tau$ lies in $D(p)$, we add the edge pq to H, and we proceed to the next pair of neighboring cells. Since $|N(\sigma)|$ is constant, a point participates in $O(1)$ point locations, of $O(\log n)$ time each. The total running time is $O(n \log n)$.

(iv): We distinguish two kinds of crossings. First, if at least one edge of a crossing lies inside a grid cell σ, then the other edge must go between different cells of $N(\sigma)$, because T_σ is crossing-free. There are $O(1)$ such edges, so there are $O(n)$ crossings of the first kind, since H has $O(n)$ edges.

In the second kind of crossing, both edges go between different grid cells. As Ψ is constant, each edge of H can participate in at most $O(1)$ such crossings. Since there are $O(n)$ edges in H, the total number of crossings is $O(n)$. □

Making G Planar. We now describe how to turn G into a planar graph. Suppose an edge pq and an edge uv of G cross at a point x. To eliminate the crossing, we add x as a new site to the graph, and we replace pq and uv by the four new edges px, xq, ux and xv. Furthermore, if qp is an edge of G, we replace it by the two edges qx, xp, and if vu is an edge of G, we replace it by the two edges vx, xu. See Figure 4. We say that this *resolves* the crossing between p, q, u and v. Let $\widetilde{G}$ be the graph obtained by iteratively resolving all crossings in G.

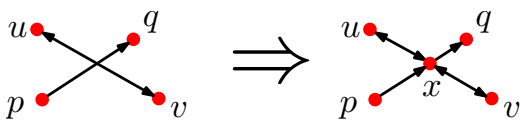

Fig. 4: Resolving a crossing. Since the edge vu exists, we also add vx and xu as edges.

First, we want to show that resolving crossings keeps the *local* reachability between the four vertices of the crossing edges. Intuitively speaking, the restriction $\Psi < 3$ forces the vertices to be close together. This guarantees the existence of additional edges between p, q, u, v in G, and these edges cover the new paths introduced by resolving the crossing.

To formally prove this, we first need a geometric observation. For a point $p \in P$, let $D(p, r)$ and $C(p, r)$ be the disk and the circle around p with radius r.

Lemma 4.4. *Let p, q be two points in $\mathbb{R}^2$ with $|pq| = \sqrt{3}$.*

(i) Let $a \in C(p, 1) \cap C(q, 1)$. Then, for any $r \in [1, \sqrt{3})$, if $b \in C(p, r) \cap C(q, r)$ lies on the other side of the line through p and q from a, then $|ab| \geq r$.

(ii) Let $\{a, b\} = C(p, \sqrt{3}) \cap C(q, 1)$. Then, $|ab| > \sqrt{3}$.

Proof. (i): Let x be the intersection point of the line segments $\overline{pq}$ and $\overline{ab}$. Then $|ab| = |ax| + |xb|$. Using that $|pa| = 1$ and $|px| = \sqrt{3}/2$, the Pythagorean Theorem gives $|xa| = 1/2$. Similarly, we can compute $|xb|$ as a function of r: with $|pb| = r$ we get $|xb| = \sqrt{r^2 - 3/4}$. We want to show that

$$r \leq |ab| = 1/2 + \sqrt{r^2 - 3/4} \Leftrightarrow r^2 \leq 1/4 + \sqrt{r^2 - 3/4} + r^2 - 3/4 \Leftrightarrow 1 \leq r^2,$$

which holds since $r \in [1, \sqrt{3})$.

(ii): Use the Pythagorean Theorem with the right angles marked in the Figure 5b. □

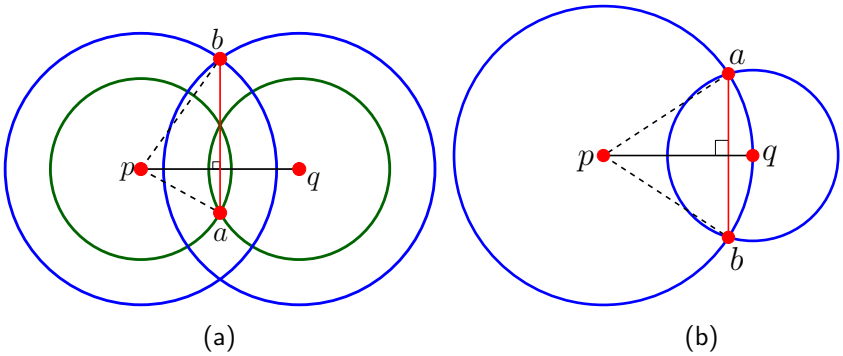

Fig. 5: The cases (i) and (ii) of Lemma 4.4.

Lemma 4.5. *Suppose that pq and uv are edges in G that cross. Let $G' \subseteq G$ be the transmission graph induced by p, q, u and v. If $\Psi < \sqrt{3}$, then p reaches v in G' and u reaches q in G'.*

Proof. We may assume that $r_p \geq r_u$. Furthermore, we set $r_q = r_v = 1$. This does not add new edges and thus reachability in the new graph implies reachability in G'. We show that if either u does not reach q (case 1) or p does not reach v (case 2), then $|uv| > r_u$. Hence uv cannot be an edge of G' despite our assumption.

Case 1: u does not reach q. Then we have $p \notin D(u)$, $q \notin D(u)$, $p \notin D(v)$ and $q \notin D(v)$. Equivalently this gives $u \notin D(p, r_u) \cup D(q, r_u)$ and $v \notin D(p, 1) \cup D(q, 1)$. Thus, the positions of u and v that minimize $|uv|$ are the intersections $u \in C(p, r_u) \cap C(q, r_u)$ and $v \in C(p, 1) \cap C(q, 1)$ on different sides of the line through p and q. To further minimize $|uv|$, observe that $|uv|$ depends on the distance of p and q and that $|uv|$ strictly decreases as $|pq|$ grows, i.e., as $|pq|$ approaches $\sqrt{3}$. For the limit case $|pq| = \sqrt{3}$, we are in the situation of Lemma 4.4(i) with $a = u$ and $b = v$ and thus we would get $|uv| \geq r_u$. But since $\Psi < \sqrt{3}$, we must have $|pq| < \sqrt{3}$ and by strict monotonicity, it follows that $|uv| > r_u$, as desired.

Case 2: p does not reach v. Then we have $u \notin D(p)$, $v \notin D(p)$, $u \notin D(q)$ and $v \notin D(q)$. We scale everything, such that $r_p = \sqrt{3}$, and we reduce r_v, r_q once again to 1. Now, the positions of u and v minimizing $|uv|$ are $\{u, v\} = C(p, \sqrt{3}) \cap C(q, 1)$. As above, further minimizing $|uv|$ gives $|pq| = \sqrt{3}$. By Lemma 4.4(ii), we have $|uv| > \sqrt{3}$ and thus uv cannot be an edge of G' (note that even after scaling we have $r_u \leq \sqrt{3}$). □

We iteratively resolve crossings in G. Call the resulting graph $\widetilde{G}$. Next, we show that for any $p, q \in P$, if p can reach q in $\widetilde{G}$, then p can also reach q in G. This seems to be a bit more difficult than one might expect, because when resolving the crossings, we introduce new vertices and edges to which Lemma 4.5 is not directly applicable.

Lemma 4.6. *For any two sites $p, q \in P$, if p can reach q in $\widetilde{G}$ then p can reach q in G.*

Proof. Each edge e of $\widetilde{G}$ lies on an edge of G with the same direction. We call it the *supporting edge* of e. A pair $p, q \in P$ such that p can reach q in $\widetilde{G}$, but not in G is called a *bad pair*. Among all bad pairs, we pick p, q such that there is a path π in $\widetilde{G}$ from p to q with the minimum number of *support switches*, where π changes from one supporting edge to another. Let $p_1q_1, \ldots, p_kq_k$ be the sequence of supporting edges as they are visited along π ($p_1 = p$, $q_k = q$).

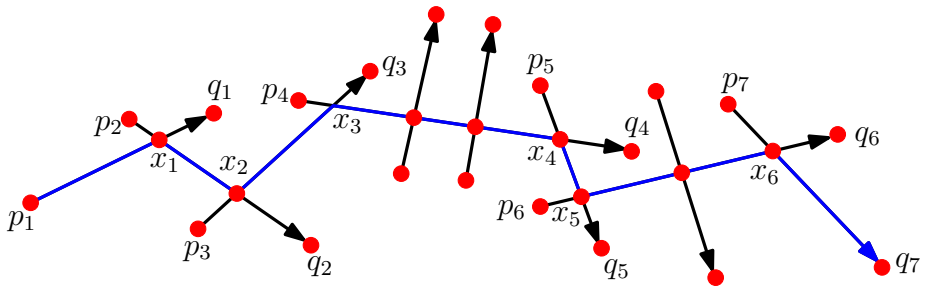

Fig. 6: A path (blue) with $k = 7$ supporting edges that is in $\widetilde{G}$ but not in G.

Claim 4.7. *The following holds in G: **(P1)** p_1 reaches $q_2, \ldots, q_{k-1}$; **(P2)** $p_2, \ldots, p_k$ reach q_k; **(P3)** p_1 does not reach $p_2, \ldots, p_k$; and **(P4)** there is no edge q_ip_i, for $i \geq 2$. Furthermore, for $i = 1, \ldots, k-1$, we have **(P5)** the line segments $\overline{p_iq_i}$ and $\overline{p_{i+1}q_{i+1}}$ have a common intersection point x_i in their interior; and **(P6)** x_{i+1} lies on the line segment $\overline{x_iq_{i+1}}$.*

Proof. **P1** and **P2** follow from the minimality of π, and **P3** follows from **P2**. For **P4**, assume that G contains an edge q_ip_i, for $i \geq 2$. By **P1**, p_1 reaches q_i in G and thus p_1 reaches p_i, despite **P3**. For **P5**, since p_iq_i and $p_{i+1}q_{i+1}$ are consecutive along π, they must intersect in a point x_i. If x_i is not in their interior, then $x_i = q_i = p_{i+1}$. But then, by **P1**, p_1 reaches $q_i = p_{i+1}$, despite **P3**. **P6** follows because by **P4** all supporting edges are directed and because the resolution of the crossings preserves the direction of the edges. □

By Lemma 4.5, we have $k \geq 3$. We now argue that the path π cannot exist. Since p_1q_1 and p_2q_2 cross, the proof of Lemma 4.5 shows that G contains one of p_1p_2, q_1p_2, p_1q_2, or q_1q_2. By **P3**, neither p_1p_2 nor q_1p_2 exist. There are two cases, depending on whether G contains p_1q_2, or q_1q_2 (see Fig. 7). Each case will lead to a contradiction with the minimality of π. s

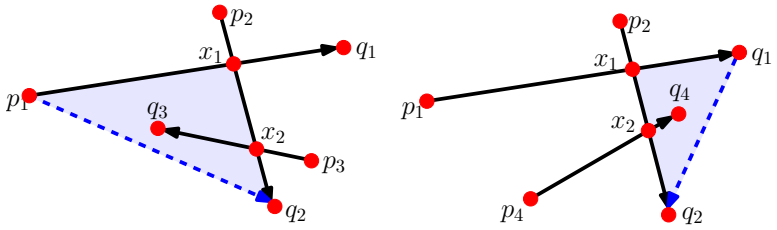

Fig. 7: Either (p_1q_2) or (p_1q_2) locks all edges in the corresponding triangle.

Case 1. G contains p_1q_2. Consider the triangle $\triangle = p_1x_1q_2$. Since $q_2, x_1 \in D(p_1)$, we have $\triangle \subset D(p_1)$. Thus, by **P3**, none of $p_2, \ldots, p_k$ may lie inside $\triangle$. By **P6**, p_3q_3 intersects the boundary of $\triangle$ in the line segment $\overline{x_1q_2}$. First, suppose that $k = 3$. In this case, we have $p_3, q_3 \notin \triangle$ (otherwise p_1 could reach q_3). Thus, p_3q_3 intersects the boundary of $\triangle$ twice, so p_3q_3 either intersects p_1q_1 or p_1q_2. In both cases, Lemma 4.5 shows that p_1 reaches q_3. Thus, we must have $k \geq 4$.

We now claim that the intersection x_3 of p_3q_3 and p_4q_4 lies in $\triangle$: if p_3q_3 intersects $\triangle$ once, then $q_3 \in \triangle$, as we already observed that $p_3 \notin \triangle$. **P6** then gives $x_3 \in \triangle$. Now suppose p_3q_3 intersects $\triangle$ twice, and let y be the second intersection point. We claim that y comes after x_2 along p_3q_3: otherwise, since x_3 comes after x_2 on p_3q_3 by **P6**, we can construct a path with fewer support switches than π: if $y \in \overline{p_1x_1}$, we omit p_2q_2; if $y \in p_1q_2$, we omit p_2q_2 and substitute p_1q_1 with p_1q_2. By the same argument, x_3 cannot come after y on p_3q_3. Thus, x_3 lies on the line segment $\overline{x_2y} \subset \triangle$. This establishes $x_3 \in \triangle$ for both cases. Now, consider the segment $\overline{p_4x_3}$. Since we observed $p_4 \notin \triangle$, we have that $\overline{p_4x_3}$ intersects $\triangle$, and we can again reroute the π to have fewer support switches.

Case 2. G contains q_1q_2. Consider the triangle $\triangle = x_1q_1q_2$. We claim that $\triangle \subset D(p_1) \cup D(q_1)$. Then the remaining argument is analogous to Case 1. Let $D(x_1) \subseteq D(p_1)$ be the disk with center x_1 and q_1 on its boundary. Let $C(x_1)$ be the boundary of $D(x_1)$. If $x_1 \in D(q_1)$, we are done. Otherwise, since the two rays from x_1 through $C(x_1) \cap C(q_1)$ intersect $D(q_1)$ inside $D(x_1)$, all line segments from x_1 to a point on $C(q_1)$ lie in $D(x_1) \cup D(q_1)$. The claim follows. □

Putting it together. To prove Lemma 4.1, we first construct the sparse subgraph H as in Lemma 4.3 in time $O(n \log n)$. Then we iteratively resolve the crossings in H to obtain $\widetilde{H}$. Since H has $O(n)$ crossings that can be found in $O(n)$ time, this takes $O(n)$ time.

Let $p, q \in P$. We must argue that p can reach q in G if and only if p can reach q in $\widetilde{H}$. Let $\widetilde{G}$ be the graph obtained by resolving the crossings in G, as in Lemma 4.6. We know that the reachability between p and q is the same in G, H, and $\widetilde{G}$. Furthermore, if p can reach q in H, then also in $\widetilde{H}$, and if p can reach q in $\widetilde{H}$, then also in $\widetilde{G}$, because (a subdivision of) every edge of $\widetilde{H}$ is present in $\widetilde{G}$. Thus, $\widetilde{H}$ and G have the same reachability properties.

4.2 Polynomial Dependence on Ψ

We now present a standard reachability oracle whose performance parameters depend polynomially on the radius ratio Ψ. Together with Theorem 2.1 we will obtain the following result:

Theorem 4.8. *Let G be the transmission graph for a set $P \subset \mathbb{R}^2$ of n points. We can construct a geometric reachability oracle for G with $S(n) = O(\Psi^5 n^{3/2})$ and $Q(n) = O(\Psi^3 \sqrt{n})$ in time $O(\Psi^5 n^{3/2})$.*

Our approach is based on a geometric separator theorem for planar disks. Let $\mathcal{D}$ be the disks induced by P. We write $\bigcup \mathcal{D} := \bigcup_{D \in \mathcal{D}} D$ and we let $\mu(\mathcal{D})$ be the area occupied by $\bigcup \mathcal{D}$. Alber and Fiala show how to find a separator for $\mathcal{D}$ with respect to $\mu(\cdot)$ [1].

Theorem 4.9 (Theorem 4.12 in [1])**.** *There exist positive constants $\alpha < 1$ and β such that the following holds: let $\mathcal{D}$ be a set of n disks and Ψ the ratio of the largest and the smallest radius in $\mathcal{D}$. Then we can find in time $O(\Psi^2 n)$ a partition $\mathcal{A} \cup \mathcal{B} \cup \mathcal{S}$ of $\mathcal{D}$ satisfying (i) $\bigcup \mathcal{A} \cap \bigcup \mathcal{B} = \emptyset$, (ii) $\mu(\mathcal{S}) \leq \beta \Psi^2 \sqrt{\mu(\mathcal{D})}$ and (iii) $\mu(\mathcal{A}), \mu(\mathcal{B}) \leq \alpha \mu(\mathcal{D})$.*

Since any directed path in G lies completely in $\bigcup \mathcal{D}$, any path from a vertex in $\mathcal{A}$ to a vertex in $\mathcal{B}$ needs to use at least one vertex of $\mathcal{S}$, see Figure 8. Since $\mu(\mathcal{S})$ is small, we can approximate $\bigcup \mathcal{S}$ with few grid cells. We choose the diameter of the cells small enough such that all vertices in one cell form a clique and are equivalent in terms of reachability. We can thus pick one vertex per cell and store the reachability information for it. Applying this idea recursively gives a separator tree that lets us answer reachability queries. Details follow.

Preprocessing Algorithm and Space Requirement. For the preprocessing phase, consider the grid $\mathcal{Q} = \mathcal{Q}_0$ whose cells have diameter 1. All vertices in a single cell form a clique in G, so it suffices to determine the reachability for one such vertex. For each non-empty cell $\sigma \in \mathcal{Q}$, we pick an arbitrary vertex $p_\sigma \in P \cap \sigma$ as the *representative* of σ. Let $R_{\mathcal{D}}$ be the set of all representatives. We recursively create a separator tree T that contains all the required reachability information: we compute $\mathcal{A}, \mathcal{B}$, and $\mathcal{S}$ according to Theorem 4.9, and we create a node v of the separator tree. Let Q_v be all cells in $\mathcal{Q}$ that intersect $\bigcup \mathcal{S}$. Let R_v be their representatives, and let $\mathcal{D}_v$ be all disks with centers in

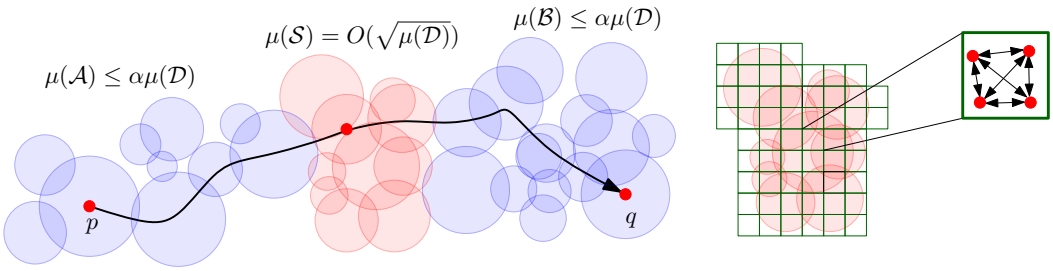

Fig. 8: Any path from $\mathcal{A}$ to $\mathcal{B}$ needs to use at least one vertex of $\mathcal{S}$. Since $\mu(\mathcal{S})$ is small, we can approximate $\bigcup\mathcal{S}$ with few grid cells.

Q_v. For each $r \in R_v$, we store all the representatives of $R_\mathcal{D}$ that r can reach and all the representatives that can reach r in the transmission graph induced by $\mathcal{D}$ (this graph is a subgraph of G). We recursively compute separator trees for $\mathcal{A} \setminus \mathcal{D}_v$ and $\mathcal{B} \setminus \mathcal{D}_v$, and we connect them to v.

To obtain the reachability information, we compute a 2-spanner H_v for the transmission graph induced by v, as in Theorem 2.2. Since we are only interested in the reachability properties of the spanner, $\varepsilon = 2$ (or any constant) suffices. For each $r \in R_v$, we compute a BFS tree in H_v with root r. Next, we reverse all edges in H_v, and we again compute BFS-trees for all $r \in R_v$ in the transposed graph. This gives the required reachability information for v.

As T has $O(\log n)$ levels, the total running time for computing the spanners is $O(n \log n(\log n + \log \Psi))$. Since the spanners are sparse, the time for computing the BFS-trees is proportional to the total number of representatives for all nodes in the tree. Below, we will show that this is at most $O(\Psi^5 n^{3/2})$. The total preprocessing time is $O(n \log^2 n + n \log \Psi + \Psi^5 n^{3/2}) = O(\Psi^5 n^{3/2})$.

To bound the space requirement, we show that $|R_\mathcal{D}| = O(\mu(\mathcal{D}))$ for any set $\mathcal{D}$ of disks.

Lemma 4.10. *Let $\mathcal{D}$ be a set of n disks with radius at least 1. Then the number of cells in $\mathcal{Q}_0$ that intersect $\bigcup\mathcal{D}$ is $O(\mu(\mathcal{D}))$.*

Proof. Suppose that a cell $\sigma \in \mathcal{Q}_0$ intersects a disk $D \in \mathcal{D}$. Then D contains a disk of radius 1 that intersects the boundary of σ. Thus, the intersection of $\bigcup\mathcal{D}$ and the region consisting of σ and its eight surrounding cells has area at least 1. Since there can be only $O(\mu(\mathcal{D}))$ different regions of this kind, the claim follows. □

Theorem 4.9(ii) and Lemma 4.10 imply that $|R_v| = O(\Psi^2\sqrt{\mu(\mathcal{D})})$ and $|R_\mathcal{D}| = O(\mu(\mathcal{D}))$, so the size of the reachability table at node v is $O(\Psi^2\mu(\mathcal{D})^{3/2})$. Thus, we obtain the following recursion for the space requirement $S(\mu(\mathcal{D}))$ for a set of disks $\mathcal{D}$ with respect to $\mu(\cdot)$:

$$S(\mu(\mathcal{D})) = S(\mu(\mathcal{A} \setminus \mathcal{D}_v)) + S(\mu(\mathcal{B} \setminus \mathcal{D}_v)) + O(\Psi^2\mu(\mathcal{D})^{3/2}). \quad (1)$$

Since Theorem 4.9 gives $\mu(\mathcal{A}) + \mu(B) \le \mu(D)$ and $\max\{\mu(\mathcal{A}), \mu(\mathcal{B})\} \le (1-\alpha)\mu(\mathcal{D})$, (1) solves to $S(\mu(\mathcal{D})) = O(\Psi^2\mu(\mathcal{D})^{3/2})$. As $\mu(\mathcal{D}) = O(n\Psi^2)$, the total space is $O(\Psi^5 n^{3/2})$.

Query Algorithm. Let $p, q \in P$ be given. We may assume that p and q are representatives for their cells. If $p = q$, then we return YES, since all vertices in the same cell constitute a clique. If $p \neq q$, we let v and w be the nodes in T with $p \in R_v$ and $q \in R_w$. Let u be least common ancestor of v and w. It can be found in $O(\log n)$ time by walking up the tree. Let L be the path from u to the root of T. We check for each $r \in \bigcup_{x \in L} R_x$ whether p can reach r and whether r can reach q. If so, we return YES. If there is no such s, we return NO. Since $|R_x|$ increases geometrically along L, the running time is dominated by the time for processing the root, which is $O(\Psi^2 \mu(\mathcal{D})^{1/2})$. Bounding $\mu(\mathcal{D})$ by $O(\Psi^2 n)$, the total query time is $O(\Psi^3 \sqrt{n})$.

It remains to argue that our query algorithm is correct. By construction, it follows that we return YES only if there is a path from p to q. Now, suppose there is a path π in G from p to q, where p and q are representatives with $p \neq q$. Let v, w be the nodes in T with $p \in R_v$ and $q \in R_w$. Let u be their least common ancestor, and L be the path from u to the root. By construction, $\bigcup_{x \in L} \mathcal{D}_x$ contains a disk for a vertex r in π. We pick r such that the corresponding node x is closest to the root. Let r' be the representative for the cell σ containing r. Since the vertices in σ constitute a clique, p can reach r' and r' can reach q in the subgraph of G induced by v. Thus, when walking along L, the algorithm will discover r' and the path from p to q. Theorem 4.8 now follows.

4.3 Logarithmic Dependence on Ψ

Finally, we improve the dependence on Ψ to be logarithmic, at the cost of a slight increase of the exponent for n. We can show the following theorem by constructing a standard reachability oracle and then using Theorem 2.1.

Theorem 4.11. *Let G be the transmission graph for a set P of n points in the plane. We can construct a geometric reachability oracle for G with $S(n) = O(n^{5/3} \log^{1/3} \Psi \log^{2/3} n)$ and $Q(n) = O(n^{2/3} \log^{1/3} \Psi \log^{2/3} n)$. All queries are answered correctly with high probability. The preprocessing time is $O(n^{5/3}(\log \Psi + \log n) \log^{1/3} \Psi \log^{2/3} n)$.*

We scale everything such that the smallest radius in P is 1. Our approach is as follows: let $p, q \in P$. If there is a p-q-path with "many" vertices, we detect this by taking a large enough random sample $S \subseteq P$ and by storing the reachability information for every vertex in S. If there is a path from p to q with "few" vertices, then p must be "close" to q, where "closeness" is defined relative to the largest radius along the path. The radii from P can lie in $O(\log \Psi)$ different scales, and for each scale we store local information to find such a "short" path.

Long Paths. Let $0 < \alpha < 1$ be a parameter to be determined later. First, we show that a random sample can be used to detect paths with many vertices.

Lemma 4.12. *We can sample a set $S \subset P$ of size $O(n^\alpha \log n)$ such that the following holds with high probability at least $1 - 1/n$: for any $p, q \in P$, if there is a path π from p to q in G with at least $n^{1-\alpha}$ vertices, then $\pi \cap S \neq \emptyset$.*

Proof. Let $m = 4n^\alpha \ln n$. We construct S by including each $p \in P$ independently with probability m/n. Using Chernoff bounds, we get

$$\Pr[|S| \geq 8n^\alpha \ln n] \leq e^{-n^\alpha \ln n} \leq \frac{1}{2n}.$$

Thus, S has size $O(n^\alpha \log n)$ with probability at least $1 - 1/2n$. Now fix p and q and let π be a path from p to q with $k \geq n^{1-\alpha}$ vertices. The probability that S contains no vertex from π is at most $(1 - m/n)^k \leq e^{-mk/n} \leq 1/n^4$, by our choice of m. Since there are $n(n-1)$ ordered vertex pairs, the union bound shows that the probability that S fails to detect a pair of vertices connected by a long path is at most $n(n-1)/n^4 \leq 1/2n$. The lemma follows by a union bound. □

We compute a sample S as in Lemma 4.12, and for each $s \in S$, we store two Boolean arrays that indicate for each $p \in P$ whether p can reach s and whether s can reach p. This needs space $O(n^{1+\alpha} \log n)$. It remains to deal with vertices that are connected by a path with less than $n^{1-\alpha}$ vertices.

Short Paths. Let $L = \lceil \log \Psi \rceil$. We consider the L grids $\mathcal{Q}_0, \ldots, \mathcal{Q}_L$ (recall that the cells in $\mathcal{Q}_i$ have diameter 2^i). For each cell $\sigma \in \mathcal{Q}_i$, let $R_\sigma \subseteq P$ be the vertices $p \in P \cap \sigma$ with $r_p \in [2^i, 2^{i+1})$. The set R_σ forms a clique in G, and for each $p \in R_\sigma$, the disk $D(p)$ contains the cell σ. The *neighborhood* $N(\sigma)$ of σ is defined as the set of all cells in $\mathcal{Q}_i$ that have distance at most $2^{i+1}n^{1-\alpha}$ from σ. We have $|N(\sigma)| = O(n^{2-2\alpha})$. Let $P_\sigma \subseteq P$ be the vertices that lie in cells of $N(\sigma)$. For every $i = 0, \ldots, L$ and for every $\sigma \in \mathcal{Q}_i$ with $R_\sigma \neq \emptyset$, we fix an arbitrary *representative point* $r_\sigma \in R_\sigma$. For every vertex $p \in P$, we store for every $i \in \{0, \ldots, L\}$ two sorted lists of cells $\sigma \in \mathcal{Q}_i$ with $p \in P_\sigma$: the first list contains all all corresponding representatives r_σ that can be reached from p; the second list contains all corresponding representatives r_σ that can reach p. A vertex p appears in at most $O(n^{2-2\alpha} \log \Psi)$ point sets P_σ, so the total space is $O(n^{3-2\alpha} \log \Psi)$.

Performing a Query. Let $p, q \in P$ be given. To decide whether p can reach q, we first check the Boolean tables for all $O(n^\alpha \log n)$ points in S. If there is an $s \in S$ such that p reaches s and s reaches q, we return YES. If not, for $i \in \{0, \ldots, L\}$, we consider the list of representatives that are reachable from p in the neighborhood at level i and the list of representatives that can reach q in the neighborhood at level i. We check whether these lists contain a common element. Since the lists are sorted, this can be done in time linear in their size. If we find a common representative at for some i, we return YES. Otherwise, we return NO.

We now prove the correctness of the query algorithm. First note that we return YES, only if there is a path from p to q. Now suppose that there is a path π from p to q. If π has at least $n^{1-\alpha}$ vertices, then by Lemma 4.12, the sample S hits π with probability at least $1 - 1/n$, and the algorithm returns YES. If π has less than $n^{1-\alpha}$ vertices, let r be the vertex of π with the largest radius, and let i be such that the radius of r lies in $[2^i, 2^{i+1})$. Let σ be the cell of $\mathcal{Q}_i$ that contains r. Since π has at most $n^{1-\alpha}$ vertices,

and since each edge of π has length at most 2^{i+1}, the path π lies entirely in the cells of $N(\sigma)$. In particular, both p and q are contained in cells of $N(\sigma)$. Since $r \in R_\sigma$ and since R_σ forms a clique in G, the representative point r_σ of σ can be reached from p and can reach q. By the symmetry of the neighborhood definition, r_σ is contained in the list of reachable representatives from p and in the lists of representatives that can reach q. This is detected when checking the corresponding lists for p and q at level i.

Time and Space Requirements. For long paths we need $O(n^\alpha \log n)$ time: for every $s \in S$ we test in $O(1)$ time whether p can reach s and whether s can reach q. For short paths there are $O(\log \Psi)$ levels, and at each level we step through two lists of size $O(n^{2-2\alpha})$. Thus, the tradeoff-point is achieved for

$$n^\alpha \log n = n^{2-2\alpha} \log \Psi \Leftrightarrow n^\alpha = n^{2/3} (\log \Psi / \log n)^{1/3}.$$

This yields $Q(n) = O(n^{2/3} \log^{1/3} \Psi \log^{2/3} n)$. This choice of α gives a space bound of $O(n^{5/3} \log^{1/3} \Psi \log^{2/3} n)$.

For the preprocessing algorithm, we first compute the reachability arrays for each $s \in S$. To do so, we build a 2-spanner H for G as in Theorem 2.2 in time $O(n(\log n + \log \Psi))$. Then, for each $s \in S$ we perform a BFS search in H and its transposed graph. This gives all vertices that s can reach and that can be reached by s in $O(n^{5/3} \log^{1/3} \Psi \log^{2/3} n)$ total time. For the short paths, the preprocessing algorithm goes as follows: For each $i = 0, \ldots, L$ and for each cell $\sigma \in \mathcal{Q}_i$ that has a representative r_σ, we compute a 2-spanner H_σ as in Theorem 2.2 for P_σ. For each representative r_σ, we do a BFS search in H_σ and the transposed graph, each starting from r_σ. This gives all $p \in P_\sigma$ that can reach r_σ and that are reachable from r_σ. The running time is dominated by the time for constructing the spanners. Since each point $p \in P$ is contained in $O(n^{2-2\alpha} \log \Psi) = O(n^{2/3} \log^{1/3} \Psi \log^{2/3} n)$ different P_σ, and since constructing H_σ takes $O(|P_\sigma|(\log \Psi + \log |P_\sigma|))$ time, the preprocessing time follows.

5 Conclusion

Transmission graphs constitute a natural class of directed graphs for which non-trivial reachability oracles can be constructed. As mentioned in the introduction, it seems to be a very challenging open problem to obtain similar results for general directed graphs. We believe that our results only scratch the surface of the possibilities offered by transmission graphs, and several interesting open problems remain.

All our results depend on the radius ratio Ψ and the major question is whether this dependency can be avoided. Our most efficient reachability oracle is for $\Psi < \sqrt{3}$. In this case the reachability of a transmission graph with n vertices can be represented by a planar graph with $O(n)$ vertices. However, it is not clear to us that the bound of $\sqrt{3}$ is tight. Can we obtain a similar result for, say, $\Psi = 100$? Or is there even a way to represent *any* transmission graph, regardless of Ψ, by a planar graph with $o(n^2)$ vertices? This would immediately imply a non-trivial reachability oracles for all ranges of Ψ.

Conversely, it would be interesting to see if we can represent the reachability of arbitrary directed graphs using transmission graphs. If this is possible, the relevant questions are how many vertices the transmission must have, what the required radius ratio is, and how fast it can be computed. A representation that achieves both few vertices and low radius ratio would lead to efficient reachability oracles for general directed graphs.

Acknowledgements. We like to thank Günter Rote for valuable comments.

References

[1] J. Alber and J. Fiala. Geometric separation and exact solutions for the parameterized independent set problem on disk graphs. *J. Algorithms*, 52(2):134–151, 2004.

[2] S. Arikati, D. Z. Chen, L. P. Chew, G. Das, M. Smid, and C. D. Zaroliagis. Planar spanners and approximate shortest path queries among obstacles in the plane. In *Proc.4thAnnu. European Sympos. Algorithms (ESA)*, pages 514–528, 1996.

[3] D. Z. Chen and J. Xu. Shortest path queries in planar graphs. In F. F. Yao and E. M. Luks, editors, *Proc.32ndAnnu. ACM Sympos. Theory Comput. (STOC)*, pages 469–478, 2000.

[4] T. Cormen, C. Leiserson, R. Rivest, and C. Stein. *Introduction to Algorithms*. 2nd edition, 2001.

[5] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, 2nd edition, 2000.

[6] H. N. Djidjev. Efficient Algorithms for Shortest Path Queries in Planar Digraphs. In *Proc.22ndInternational Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, pages 151–165, 1996.

[7] G. N. Federickson. Fast algorithms for shortest paths in planar graphs, with applications. *SIAM J. Comput.*, 16(6):1004–1022, 1987.

[8] J. Holm, E. Rotenberg, and M. Thorup. Planar Reachability in Linear Space and Constant Time. *CoRR*, arXiv:1411.5867, 2014.

[9] H. Kaplan, W. Mulzer, L. Roditty, and P. Seiferth. Spanners for Directed Transmission Graphs. *CoRR*, arXiv:BLA ADD MEE, 2015.

[10] D. Peleg and L. Roditty. Localized spanner construction for ad hoc networks with variable transmission range. *ACM Transactions on Sensor Networks (TOSN)*, 7(3), 2010.

[11] M. Pătraşcu. Unifying the Landscape of Cell-Probe Lower Bounds. *SIAM J. Comput.*, 40(3):827–847, 2011.

[12] M. Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *J. ACM*, 51(6):993–1024, 2004.

[13] P. von Rickenbach, R. Wattenhofer, and A. Zollinger. Algorithmic Models of Interference in Wireless Ad Hoc and Sensor Networks. *Networking, IEEE/ACM Transactions on*, 17(1):172–185, 2009.